

AD640962

MEMORANDUM

RM-5143-PR

SEPTEMBER 1966

A DEEP INDEX OF
DERIVATIONAL MORPHOLOGY

Dean S. Worth

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION			
Hardcopy	Microfilm		
\$2.00	\$.50	35	PP 72
1 ARCHIVE COPY			

OCT 31 1966
B

PREPARED FOR:

UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

BEST AVAILABLE COPY

20040910154

MEMORANDUM

RM-5143-PR

SEPTEMBER 1966

A DEEP INDEX OF
DERIVATIONAL MORPHOLOGY

Dean S. Worth

This research is sponsored by the United States Air Force under Project RAND—Contract No. AF 49(633)-1700—monitored by the Directorate of Operational Requirements and Development Plans, Deputy Chief of Staff, Research and Development, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.

DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

The RAND Corporation

1700 MAIN ST. • SANTA MONICA • CALIFORNIA • 90406

PREFACE

The present Memorandum brings together two tasks being performed at RAND. A derivational dictionary of Russian is being constructed as the basis for a series of studies of affixation, deep-structure morphophonemics, etc. To facilitate access to the extant literature in the field, a deep index is being prepared with the assistance of the "Catalog" system.*

The present Memorandum, which is concerned with classification and condensation of information dealing with substantive problems of Russian morphology, will be of interest principally to students of the language, including those in the Air Force who are using or developing systems for automatic processing of the language.

* Martin Kay and Theodore Zieve, The Catalog: A Flexible Data Structure for Magnetic Tape, The RAND Corporation, RM-4645-PR, October 1965.

SUMMARY

This Memorandum describes a projected "Deep Index" to the extant literature on Russian derivational morphology. The Memorandum describes the type of literature being abstracted, the classification of information into discrete categories for retrieval, the steps taken to prepare data for conversion to machine format, and the transliteration and abbreviation conventions used in the Index. An example of Russian scholarly text is appended, together with the format in which this text appears after processing.

ACKNOWLEDGMENTS

Both the general goals of the Deep Index project, and the details of abstracting and conversion to machine format have been discussed at length with D. G. Hays, T. Ziehe, and H. Noguni, whose helpful suggestions are gratefully acknowledged.

A DEEP INDEX OF DERIVATIONAL MORPHOLOGY

1. INTRODUCTION

In what follows we describe a project, called Deep Index, established to gather, file, and retrieve a large quantity of information concerning the derivational morphology of Russian. This work is part of a larger, ongoing study of the morphological structure of Russian, during which some 110,000 Russian words have been put on magnetic tape, segmented into their component morphemes, and organized into derivational families or wordnets.* The unusual size of the corpus available for analysis, together with the possibility of utilizing sophisticated computational techniques in this analysis, holds out the prospect of novel and interesting scientific conclusions. Such conclusions, however, must be matched against the substantial body of fact and opinion that has accumulated during many decades of more traditional study. The sterile rediscovery of known facts can be avoided only by thorough control of the relevant literature; the Deep Index project has been established to provide such control.

The first months of the Deep Index project were spent in devising and improving the format in which information from the literature was to be abstracted and filed, and with determining the optimal procedures for transferring abstracted

* The 110,000-word derivational dictionary ("Word-Family Glossary"), prepared by D. S. Worth and A. S. Kozak, is described in more detail in a forthcoming RAND Memorandum. The dictionary itself will be ready for publication in early 1967.

material in English abbreviations onto magnetic tape. These preliminary tasks, which will be described in more detail below, have now been completed; during the next stage of the project, the principal effort expended will be to abstract and put into machine format as much relevant information as possible. As the file of data grows, it will become possible to ask for retrieval of increasingly valuable collocations of information.

2. LITERATURE TO BE PROCESSED

The literature that has been, is being, and will be abstracted and fed into the Deep Index deals primarily with the derivational morphology of contemporary standard Russian (CSR), with priority given to more recent works and to those that utilize at least some of the attainments of structural linguistics. Since, however, the morphophonemic processes of the CSR derivational system (in particular, the phonological alternations occasioned by the concatenation of stems and affixes) are the result of phonetic changes that reach back into the distant past, one cannot neglect entirely the considerable literature dealing with the history of derivational patterns, with increases or decreases in productivity, and with the origin and development of individual affixes. Within the limits imposed by time, availability of competent staff, facilities (machine time, etc.)

and the availability of the literature itself, the Deep Index project will attempt to abstract and to add to the data file the bulk of the current literature on the topic, and at the same time, but with lesser intensity, to extend its coverage both temporally and in subject matter--that is, to include in the file at least a portion of the enormous bulk of extant literature from provincial universities in the Soviet Union, from the scholars of preceding generations, or from literature dealing with problems whose relevance to the derivational morphology of CSR is tangential rather than direct (e.g., comparisons between Russian and other Slavic languages, the influence of productive derivational patterns in nineteenth-century French on those of Russian at that time, etc.). A tempting prospect for the more distant future would be to expand the project to cover the derivational morphology of all Slavic languages in all their synchronic and diachronic variety. For the moment, however, we are concerned primarily with providing a reliable research facility for investigators of derivation in CSR itself.

3. TYPOLOGY OF INFORMATION

The information that is extracted from literature being processed deals with both items and processes. The primary orientation, both in abstracting and in retrieving, is toward affixes, partially because the overwhelming bulk of derivational

procedures involves specific affixes (although affixation by no means exhausts the subject of derivation), and partially because affixes provide more economical and less ambiguous items of information for storage and retrieval than do statements dealing with processes ("truncation", "palatalization", "interfixation", "inclusion", etc.). While we agree completely with Stankiewicz that the derivational system of a language must be treated in terms of its structural relations, that is, indeed as a system,* we consider it essential that this system be described in terms of items and their frequency, and of processes and their productivity, with both items and processes described explicitly and precisely.

Most but not all of the items of information abstracted fall into a standard pattern. This includes an identification of the specific affix under discussion, the stem type or types (noun or verb stem, toponym, etc.) to which the given affix can be added, the type of morphophonemic change occasioned by the affixation (palatalization, truncation, etc.), the meaning class of the resultant word, the grammatical features of the resultant word (feminine noun, perfective verb, etc.), such historical information as may be provided, and one or more examples of the given type of derivation. In general, we have tried to devise a system

* Stankiewicz, E., "The Interdependence of Paradigmatic and Derivational Patterns," Word, Vol. 18, 1962, p. 3.

as flexible as possible, feeling that future experience may well lead us to include different categories of information into the file, and to search for combinations of item information and process information within fixed parameters (e.g. -- a purely hypothetical example, which could become real only when a much larger data file is available -- it might be convenient to retrieve such collocations of information items as "feminine substantives in -k(a) involving stem truncation and forming toponyms in North Great Russian dialects"). While we would not anticipate any major changes in the format of the Deep Index itself, the number and type of search categories (i.e., the information items specifically classified and retrievable as members of the specified class) have been left open for such future additions as may prove desirable.

Let us examine the abstracting process in more detail. When relevant information is encountered in text, it is classified according to the affix being discussed. Page number in text is stated, the affix is listed, and the information pertaining to this affix is placed in abbreviated form in a specific format (see below) after the affix. In cases where the information being recorded deals with more than one affix, it is repeated for each individual affix; this extra task in preparing the information for filing is repaid in ease of retrieval. The information

thus provided for each affix is broken down into one or more of the following categories (given with their abbreviations):

Affxd = Affixed. This category describes the morphotactics of the affix, i.e., with which stem types the given affix appears. E.g.: enepthetic hs is added to indeclinable noun stems before further affixes are added (kino → kino-hs-nyj, kino-hs-ka).*

Forms = Information in this category describes the product of the given derivation, the result of the particular affixation procedure described, e.g., -k(a) forms feminine (occasionally animate) nouns from adjective + noun combinations (texnihceskaha sluhza-haha → texnihcka).

Prod = Productivity--whether (and if so, to what degree) an affix is used to form new words in CSR.

Dist = Distribution. This category contains information concerning the morphophonemics of the affix, i.e., the distribution of allomorphs of an affixal morpheme in terms of the stem types with which each can occur.

* For the transliteration system used in this paper, see Sec. 5.

E.g. the affix -hcanin occurs in the form -hcanin after stems not ending in hc or a consonant alternating with hc; the form -anin occurs with such stems as Tambov tambovhcanin; Kuzneck → kuznehcanin.

- Chng = Change. This category contains information concerning the morphophonemics of the stem (cf. "Dist" above), i.e., the stem alternations occurring in the presence of a given affix. E.g., adjective stems in -telqn- lose this segment and add -l- before the -k(a) that forms feminine nouns from adjective + noun combinations (zahzigatelqnaha bomba → zahzigalka).
- Mn = Meaning (of the affix, of the derived stem, or of the stem to which the affix is added). E.g., the meaning of English common-gender nouns formed with the suffix -tor is 'noun of agent' or 'implement' (operator).
- Stl = Style (of the affix or of the derived stem). E.g., derivatives in -k(a) from adjective + noun combinations are colloquial in style (Vehcernhaha Moskva → vehcerka).
- Hist = Historical information concerning the given affix or a process connected with it.

- Rel = Relation of the given affix to other affixes (whether it is homonymous, synonymous, etc.). E.g., the affixes -lqn(yj) and -telqn(yj) are synonymous in relative adjectives (promyvalqnyj, promyvateľqnyj).
- Epen = Epenthetica or "interfixes" (empty morphs, Verbindungsmorphemen). This is not a category of information, but a class of items, the morphemic status of which is so unusual that it was felt appropriate to identify it specifically for ease in retrieval. E.g., the hs in kinohska, the j in hsossejnyj, the v in hemgeuvskij from MGU, etc.
- X = Example (of an affix, process, etc.). This marker is ordinarily followed by a representative example of the affix or process being described, in transliteration from the Cyrillic alphabet. Not all examples given in the original text are transcribed into the data file. (This category is discussed further below, pp. 9-10.)
- Misc = Miscellany. This category includes information that does not fit unambiguously into any of the above categories, but that may later be subdivided into additional specific categories, and for which appropriate labels

will become evident only after larger quantities of text are processed. I.e., this category is a means of getting data into the file without at this time making a firm commitment as to their classification; it is one way of building flexibility into the system.

The information contained in each of these categories forms a discrete unit, marked in the file by the fact that the categorial heading begins with a capital letter and that the category ends with a period. Every item of information that is appropriately classified under one of these categories must be labeled by the appropriate abbreviation.

Information in the data file will thus be in the form of paragraphs, each headed by a transliteration of the affix to which the information in the paragraph pertains, and each containing a number of categories of information (cf. above) any or all of which can be specified in the retrieval process. Complete bibliographical information concerning the document from which the information was abstracted is filed and will be retrieved automatically together with whatever information is asked for from the file.

In certain, undoubtedly infrequent cases, an item of information may be appropriately classified in more than one category. In such cases, the information will be filed

in the form of a single sentence, which will contain all the appropriate category abbreviations. During the retrieval process, this sentence will be printed out for one or all of the category headings, according to which type of information is requested. The most common example of this is the category "X", or "example". It is used as an independent category only where the given derivation is rare or otherwise unusual; however, examples are also given (and called "X") as part of another category-sentence to elucidate the material contained in the category.

Thus if examples of a given affix are desired, they will be printed out whether "X" is contained under another heading or is independent. Should "X" not be independent the whole sentence in which "X" is contained will be printed out, i.e., the whole category will be printed out. Likewise, a call to print a heading that includes examples will include "X".

4. TECHNICAL DETAILS

The process of abstracting must of course proceed according to a fixed format and with fixed symbolic conventions. The abstractor must read through the article or book chapter completely before beginning to abstract the information contained therein (this for the obvious reason that the value of the entire Deep Index hinges on the intelligent classification of information by linguistically sophisticated abstractors who understand thoroughly the material with which

they are dealing). During the second reading, whenever an item of information is encountered that should be entered in the data file (e.g., whenever some suffix is described in terms of meaning, stems it can be added to, etc.), the abstractors perform the following:

(1) The page reference is recorded. Since the article or book title is given before any material is abstracted, this makes it possible to print out complete bibliographical information for any items retrieved.

(2) The affix is transliterated in square brackets. All allomorphs of the affix that are given in the source are copied out, e.g. [-sk(ij)/-ovsk(ij)]. Square brackets are used throughout the Index to provide an unambiguous signal on tape indicating the beginning and end of every Cyrillic sequence; the purpose of this signal is to make it possible to print automatically from tape in both the Cyrillic and Latin alphabets. (There are other methods of providing such signals, but we have used brackets for convenience--they appear on the preliminary typed copy and are easy to proof-read.)

(3) All information in the source relevant to the given affix is then entered in a single paragraph to the right of the affix. This information must be in the form of discrete sentences, classified into the categories discussed in Sec. 3. These paragraphs are of course the heart of the project; the value of the Index depends

on the accurate interpretation and classification of the at times diffuse and contradictory statements in the sources, and these data must be filed in a fixed format to make it possible to retrieve all instances of identically classified information. Ideally, the Deep Index project would require an elaborate taxonomy of knowledge, and one on which all authors and all abstractors were in perfect agreement; there would exist no vague statements, no terminological misunderstandings, etc. The Deep Index project is predicated upon the assumption not that such an ideal situation exists, but that it is important to try to move in that direction.

When an article or book chapter has been completely abstracted as outlined above, the material is typed on a Dura Mach-10 typewriter, which provides typed copy for proof-reading while simultaneously transcribing this copy in coded form onto punched paper tape, thus facilitating both corrections in the copy (after proof-reading) and the transcription of this copy onto magnetic tape for storage in the data file. The data are transferred and stored in the Kay-Ziehe "Catalog" format.* Since the procedures and problems of searching and print-out are not unique to this project, there is no need to discuss them here.

* Kay, Martin, and Theodore Ziehe, The Catalog: A Flexible Data Structure for Magnetic Tape, The RAND Corporation, RM-4645-PR, October 1965.

Two technical problems that had to be solved were that of transliteration from the Cyrillic into the Latin alphabet, and that of devising an appropriate set of abbreviations to render the material being transcribed and filed less bulky than if filed in unabbreviated English.

5. TRANSLITERATION

Thanks to a convenient coincidence, an ingenious system of transliteration into and out of the Cyrillic alphabet was proposed by R. Jakobson^{*} just as the Deep Index project was getting under way. Jakobson's proposals were intended to facilitate telegraphic communication in Cyrillic, but have proved extremely useful for the present project as well. The transliteration system makes use of all twenty-six letters of the Latin alphabet. Two of these are used for Cyrillic letters having no natural correspondents in Latin: q = б and w = в. The only "diacritic" used in the Jakobson system is the Latin h, which has no natural correspondent in Cyrillic and is usually unused in transliteration (except for digraphs like kh = Cyrillic x in the Library of Congress system). Latin h in Jakobson's system serves as a parameter shifter, giving the characters z, c, s and o, e, u, a the "meanings" of Cyrillic ж, ч, ш, and ё, э, ю, я instead of з, ц, с and о, е, у, а, respectively. Particularly ingenious is the use of Latin h to change its own "meaning":

^{*}Jakobson, R., Voprosy hazykoznanija, 1965, No. 1, pp. 111-113.

the sequence hh corresponds to Cyrillic н; i.e., the first h has the "meaning" 'read an immediately following h not as a parameter shifter but as Cyrillic н'. The entire system of correspondences is as follows (in the order of the Cyrillic alphabet):

а	б	в	г	д	е	ё	ж	з	и	й	к	л
a	b	v	g	d	e	ho	hz	z	i	j	k	l
м	н	о	п	р	с	т	у	ф	х	ц	ч	ш
m	n	o	p	r	s	t	u	f	x	c	hc	hs
щ	ъ	ы	ь	э	ю	я						
hh	w	y	q	he	hu	ha						

This system, although somewhat disconcerting at first glance, becomes as easy to use as any other after a very brief period.

6. ABBREVIATIONS

As a space-saving device, the following list of abbreviations was drawn up, based on terms that occurred with particular frequency during the excerpting of the first articles being entered into the Index (cf. also the categorial abbreviations on pp. 6-9 above):

A	Adjective, adjectival
Abb	Abbreviation: initial or syllabic abbreviations that function as words
Abst	Abstract, of a noun

acc	according to. . .
Act	Active, active voice
Ad	Adverb, adverbial
Aff	Affectionate
Allm	Allomorph
An	Animate
-An	Inanimate
Anal	Analogy, as in "analogical leveling"
Aug	Augmentative
Bul	Bulgarian
BR	Belorussian
c	consonant, consonantal
Col	Colloquial
Comb	Combination
CSR	Contemporary Standard Russian
Cz	Czech
Decl	Declinable, declension
-Decl	Indeclinable
Der	Derivation, derivational, derived
des	designates, designating
Desc	Descriptive, as of an A
Det	Determined aspect
-Det	Non-determined aspect
Dial	Dialectal
Dim	Diminutive
Dsyl	Disyllable, disyllabic

Eng	English
esp	especially
Exp	Expressive
Fem	Feminine
For	Foreign
Fr	French
G	Gerund
Gen	Gender
Ger	German
Gr	Greek
Hom	Homonymous, homonym
I	Infix, interfix: cf. "Epen", p. 8.
Imp	Imperfective aspect
Incl	Inclusion, i.e., a morphological process involving the derivation of a single word from a common combination of two or more words such that the meanings of all the words in the combination are <u>included</u> in the single word derived.
incl	includes, inclusive
Int	Interjection
Lat	Latin
M	Morpheme
Mac	Macedonian
Masc	Masculine
Msyl	Monosyllable, monosyllabic
N	Noun
Neut	Neuter

Ohj	Object as opposed to "person"; <u>predmet</u> , <u>predmetnoe znahcenie</u>
occ	occasionally
OCS	Old Church Slavic
Or	Origin
P	Prefix
Poid	Prefixoid
Pal	Palatal
Palzd	Palatalized
par	particularly
Pass	Passive, passive voice
Pej	Pejorative
Perf	Perfective
Plur	Plural
Pol	Polish
Pr	Pronoun
Prep	Preposition
Pt	Participle
Qual	Qualitative, as of an A.
Rl	Relative, as of an A.
Russ	Russian
S	Suffix
Soid	Suffixoid
SC	Serbo-Croatian
sel	seldom
Sem	Semantic: semantic groups, e.g., time expressions

Sing	Singular
St	Stem
Sub	Substantive
Subzd	Substantivized
Syl	Syllable, syllabic
Syn	Synonymous, synonym
Trans	Transitive
-Trans	Intransitive
Trunc	Truncated, truncation
Ukr	Ukranian
us	usually
V	Verb, verbal
v	vowel
Wd	Word
=	is, includes, etc.
--	is not, does not include, etc.
#	Derivation involving loss of affix(es) and a reversal of the morphophonemic processes that occurred along with the affix(es), e.g., <u>stolihcna</u> vodka = <u>stolica</u> .
*	Zero affixation (i.e., derivation without affixation), e.g. <u>uklonitq</u> → <u>uklon</u> .

7. CAPITALIZATION

Those terms (abbreviated or not) that mark categories of information, names of languages, parts of speech, derivational procedures, etc., for which the data file will be searched, are called "significant". Significant terms are signalled by capitalization, without regard to their position within a sentence; concomitantly, non-significant terms are never capitalized, even in sentence-initial position. Non-significant terms can of course be searched and retrieved if the occasion should require this, but it appears unlikely that this will be the case.

Capitalization also marks the difference between otherwise homographic forms: "Incl" means 'Inclusion' (Russian vkluhcenie), whereas "incl" means 'inclusive', 'includes' (non-technical terms).

8. EXAMPLE OF ABSTRACTING FORMAT

As an example of the manner in which source material is abstracted and classified, we cite below a section of text from an article on inclusion by N.A. Hanko-Trinickaha, "Processy vkluhceniha v leksike i slovoobrazovanii," which appeared in the collection Razvitie grammatiki i leksiki sovremennogo russkogo hazyka, Moscow, 1964, pp. 18-35. The passage appearing below occurs on pp. 27-29 of this article, and describes the inclusion of the meaning of a modified noun into a derivative based on the adjective stem of an original adjective + noun combination, with the suffix -k(a).

We give the passage in the Russian original, in full English translation, in the abbreviated format prepared by the abstractor, and finally in such a format with abbreviations resolved.

8.1. Russian Text

Б. Включение определяемого. При включении определяемого происходит изменение грамматического разряда слов, что выражается изменением основы слова (суффиксация) и системы флексий. Обычно это образование существительных от основы прилагательных на фоне сочетания данных прилагательных с существительными.

Наиболее типичным представителем данного явления можно считать существительные жен. рода, возникающие на базе привычных сочетаний прилагательных с существительными. В основу производного слова кладется основа прилагательного, обычно освобожденная частично или полностью от суффиксов, и присоединяется суффикс -ка в сопровождении системы флексий первого склонения, например: *одиночка* 'одиночная камера', *ротационка* 'ротационная машина', *шоссейка* 'шоссейная дорога', *вечерка* '«Вечерняя Москва» или вообще вечерняя газета', *зачетка* 'зачетная книжка', *Публичка* 'Публичная библиотека', *Таганка* 'Таганская площадь' и т. п.

Существительное, значение которого включается в производное слово, может принадлежать не только к жен. роду, но и к мужскому, и к среднему, а иногда даже к мн. числу, например: *литейка* 'литейный цех', *земледелка* 'земледельный цех', *электричка* 'электрический поезд', *короткометражка* 'короткометражный фильм', *Александринка* 'Александринский театр', *касторка* 'касторовое масло', *караулка* 'караульное помещение', *родилка* 'родильное отделение', *валерианка* 'валериановые капли' и др.

Большинство слов этого образца возникает в непринужденной разговорной речи и как всякое стилистически окрашенное новое слово не всегда и не сразу становится общим достоянием, например *безымянка* 'безымянный палец', *бертолетка* 'бертолетова соль', *психиатричка* 'психиатрическая больница'.

Можайка 'Можанское шоссе', *фугаска* 'фугасная бомба', *пятидневка* 'пятидневная неделя', *водородка* 'водородная бомба', *обручалка* 'обручальное кольцо', и т. п.

Значительная часть таких образований имеет обращение в более или менее узком кругу, где хорошо известно соответствующее словосочетание, например: *Тургеневка* 'Тургеневская библиотека', *Горьковка* 'Горьковская библиотека', *зарубежка* 'зарубежная литература' (учебный курс), *мореходка* 'мореходное училище', *дивизионка* 'дивизионная газета' и т. п. Ср.: Но в группе туристов не случайно не очень весело шутили, что у нас на «Воробьевке» такой инструктор, как Экберт, за один сезон может сделать больше, чем целая команда московских тренеров, которые учат по старинке («Комс. правда», 21 марта 1961 г.).

Это настолько живой и продуктивный способ, что его не удерживает даже возникновение омонимов как со словами других образцов, так и со словами, образованными по этому же образцу, но включающими значение другого существительного, например: *комсомолка* (девушка и «Комсомольская правда»), *пионерка* (девочка и «Пионерская правда»), *санитарка* (женщина-специалист и санитарная машина), *вегетарианка* (лицо женского пола и вегетарианская столовая), *испанка* (лицо женского пола и испанская инфлуэнца, свирепствовавшая в 1918 г.), *ленинка* (последовательница Ленина и Ленинская библиотека), *кандидатка* (женщина-кандидат и кандидатская работа), *самоходка* (самоходная пушка и самоходный комбайн), *стационарка* (стационарная школа, стационарная больница, стационарная киноустановка), *районка* (районная поликлиника, районная больница) и т. п.

Интересно, что по этому образцу с включением иногда образуются названия лиц женского пола, например: *техничка* «техническая служащая»; После «Генриады» славимой «*Орлеанку*» сочинил Волтер, повесть шуточную, вредную, но остроумную (М. М. Херасков, Бахарина, или Неизвестный). *Орлеанка* «Орлеанская дева».

Ср. появившееся в наши дни *Сикстинка* «Сикстинская мадонна»: Как бы нам количество переделать в качество? Это сделать очень просто, в плане творческого роста: в год хотя бы по картинке, но на уровне *Сикстинки* («Московский художник», 25 дек. 1958 г.); [Олег] Вы королева, вы богиня, вы *сикстинка*! [Маша] (возмущенно) Что ... что такое? [Олег] Это не ругательство, честное слово, Сикстинская мадонна! (Г. Трифонов, Обязательное свидание).

Формальное ограничение данного образца идет по линии производящих прилагательных: для него не используются прилагательные односложной основы. Поэтому не может возникнуть подобное существительное на базе, например, словосочетания *скорая помощь*, но легко возникает на базе словосочетания *неотложная помощь* — *неотложка*. В первые годы после революции было создано слово *стенка* «стенная газета», но оно не удержалось и было заменено словом *стенновка*. Ср.: Л. Г. Бальяна научно изучала воспитательную роль школьной *стенновки* («Лит. газета», 11 апр. 1961 г.).

Производящие основы прилагательных претерпевают некоторые изменения, утрачивая полностью или частично суффиксы: *моторная лодка* — *моторка*, *снайперская винтовка* — *снайперка*, *анатомический театр* — *анатомичка*, *карболовая кислота* — *карболка*, *комиссионный магазин* — *комиссионка*.

Однако суффикс, имеющий в своем составе гласный, не утрачивается, если без него основа оказывается односложной: *овсяная крупа* — *овсянка*, *ветряная оспа* — *ветрянка*, *взрывчатое вещество* — *взрывчатка*, *зубчатое звено* — *зубчатка*, *керзовые сапоги* — *керзовки*. Ср. также возникшее некогда путем включения *винтовка* (винтовое ружье).

Прилагательные с суффиксом *-тельн-* заменяют его на суффикс *-л-*: *зажигательная бомба* — *зажигалка*, *курительная комната* — *курилка*.

Пока слово с включением употребляется наряду со словосочетанием, оно имеет сниженную стилистическую окраску по сравнению с последним. Употребляясь все шире, оно оттесняет словосочетание, может становиться единственным, а потому и нейтральным обозначением данного предметного значения, и самое возникновение его на базе словосочетания забывается: *перчатка*, *косоворотка*, *землянка*, *открытка*, *пятилетка* и под.

8.2. ENGLISH TRANSLATION

B. Inclusion of the modified word: In inclusion of the word modified there occurs a change in the grammatical class of words, which is expressed by a change in stem of the word (suffixation) and in the system of inflexion. This usually involves the formation of a noun from an adjective against the background of a combination of the given adjective and a noun.

The most typical representative of this phenomenon may be considered to be nouns of feminine gender derived on the basis of ordinary combinations of adjective and noun. The stem of the adjective is taken as the stem of the derived noun, with the adjective usually losing its suffixes partially or wholly and the suffix -k- added, accompanied by the inflexional system of the first declension, for example: odinoɦka < odinoɦcnaha kamera 'solitary confinement', rotacionka < rotacionnaha mahsina 'rotary press', hsossejka < hsossejnaha doroga 'macadam road', vehcerka < Vehcernhaha Moskva 'Moscow Evening News', or in general, 'an evening newspaper', zahcetka < zahcetnaha knihzka 'record book', Publihcka < Publihcnaha biblioteka 'Public Library', Taganka < Taganskaha plohhadq 'Tagansk Square', etc.

The noun, the meaning of which is included in the derived word, can belong not only to the feminine gender but can also be masculine or neuter and sometimes even plural, for example: litejka < litejnyj cex 'foundry shop or department', zemledelka < zemledelqny cex 'agricultural department', helektrihcka < helektrihceskij poezd 'electric railway train', korotkometrahzka < korotkometrahznyj filqm 'short film', Aleksandrinka < Aleksandrinskij teatr 'Aleksandrinskij Theatre', kastorka < kastorovoe maslo 'castor oil', karaulka < karaulqnoe pomehhenie 'guard house', rodilka < rodilqnoe otdelenie 'delivery room', valerianka < valerianovye kapli 'tincture of valerian', and others.

Most words of this model occur in informal, conversational speech and like every stylistically colored new word, they are not always or immediately generally accepted, for example: bezymhanka < bezymhannyj palec 'ring finger', bertoletka < bertoletova solq 'potassium chlorate',

psixiatrihcka < psixiatrihceskaha bol'nica 'mental hospital', Mohzajka < Mohzajskoe h'sosse 'Mozajsk Highway', fugaska < fugasnaha bomba 'demolition bomb', phatidnevka < phatidnevnaa nedel'ha 'five-day week', vodorodka < vodorodnaa bomba 'hydrogen bomb', obruhcalka < obruhcalknoe kol'co 'wedding ring', etc.

A significant portion of such formations are restricted to a more or less narrow circle where the corresponding word-combination is well known, for example: Turgen'evka < Turgenevskaha biblioteka 'Turgenev Library', Gork'ovka < Gork'ovskaha biblioteka 'Gorkij Library', zarubezhka < zarubezhnaa literatura 'foreign literature, an educational course', morexodka < morexodnoe uhcilishhe 'nautical school', divizionka < divizionnaa gazeta 'division newspaper', etc. Cf.: "It was not by chance that in a group of tourists someone joked rather sadly about the fact that we have an instructor at the Vorob'evka such as Ekbert who can do more in one season than a whole team of Moscow trainers who teach the old-fashioned way." (Komsomol'skaha pravda, 21/III/1961).

This is such a live and productive method that it is not even restricted by the appearance of homonyms based both on words of different models and on words formed according to the very same model, but including the meaning of a different noun, for example: komsomolka 'girl (member of the Komsomol)' and Komsomol'skaha pravda, pionerka 'girl (member of the Pioneers)' and Pionerskaha pravda, sanitarka 'woman specialist' and 'army ambulance', vegetarianka 'female vegetarian' and 'vegetarian dining hall', ispanka 'Spanish woman' and 'Spanish influenza which raged in 1918', leninka 'female follower of Lenin' and 'Lenin Library', kandidatka 'female candidate' and 'candidate's work', samoxodka 'self-propelled artillery' and 'self-propelled combine', stacionarka 'stationary school', 'stationary hospital' or 'stationary cinema unit', rajonka 'regional polyclinic', 'regional hospital', etc.

It is interesting that this model of inclusion sometimes forms the names of female persons: texnihcka < texnihceskaha sluhzakhaha 'female technical employee'. "After his celebrated La Henriade Voltaire wrote Orleanka, a comic, pernicious, but sharp-witted story." (M.M. Xeraskov, Baxarina, or the Unknown). Orleanka < Orleanskaha deva 'Maid of Orleans'.

Cf. the recent formation Sikstinka < Sikstinskaha madonna 'Sistine Madonna'. "How can we transform quantity into quality? This is easy to do in the plan of creative growth: even if there's but one picture a year, but then -- on the level of the Sikstinka." (The Moscow Artist, 25/XII/1958); "[Oleg] You're a queen, a goddess, a sikstinka! [Mahsa] (indignantly) What... What's that....? [Oleg] It's not a swear-word, cross my heart, it's the Sikstinskaha madonna! (G. Trifonov, The Necessary Meeting).

The formal limitations of this model are set by the adjectives which serve as stems: monosyllabic adjective stems cannot be used. Therefore, a noun cannot be derived, for example, from the word-combination, skoraha pomohhq 'ambulance', but may easily be derived from neotlohznaha pomohhq > neotlohzka 'first aid'. During the first few years after the revolution the word stenka < stennaha gazeta 'wall newspaper' was coined but it did not survive and was changed to stennovka. Cf. "L.G. Balqhan made a scientific study of the educational role of stennovki." (Literaturnaha gazeta, 11/IV/1961).

The adjective stems serving as derivational bases are subjected to changes, losing their suffixes partially or wholly: motornaha lodka > motorka 'motor boat', snajperskaha vintovka > snajperka 'sniper's rifle', anatomihceskij teatr > anatomihcka 'anatomical theater', karbo-lovaha kislota > karbolka 'carbolic acid', komissionyj magazin > komissionka 'second-hand shop'.

However, a suffix containing a vowel is not truncated if the stem would become monosyllabic without it: ovshanaha krupa > ovshanka 'oatmeal', vetrhanaha ospa > vetrhanka 'chicken pox', vzryvhcatoe vehhestvo > vzryvhcatka 'blasting charge', zubhcatoe zveno > zubhcatka 'gear', kerzovye sapogi > kerzovki 'Curzon boots'. Cf. the old inclusion vintovka < vintovoe ruhze 'rifle'.

Adjectives in -telqn- replace that suffix with -l-: zahzigatelqnaha bomba > zahzigalka 'incendiary bomb', kuritelqnaha komnata > kurilka 'smoking room'.

As long as the inclusion is used side by side with the word-combination, it has a lowered stylistic coloring in comparison with the latter. Becoming more widely used, it crowds out the word-combination and can become the only (and therefore the neutral) designation of the given referential meaning and the very fact that it was derived on the basis of a word-combination is then forgotten: perhcatka 'glove', kosovorotka 'blouse which fastens at the side', zemlhanka 'dug-out, mud hut', otkrytka 'post-card', phatiletka 'Five Year Plan', et al.

8.3. Abstractors' Format

Following is the above passage as abstracted and condensed; this is the format in which excerpted material is filed in the catalogue:

[Hanko-Trinickaha, N. A., "Processy vkluhceniha v leksike i slovoobrazovanii," Razvitie grammatiki i leksiki sovremenogo russkogo hazyka], Moscow, 1964, pp. 18-35.

- 25 [-sha]: Affxd Trans V. Mn '-An Obj, reflexive, mutual, or generalized Obj'.
- 27 [-k(a)]: Affxd A-St of AN Comb; 'ncl mn of N, X, [helektrihceskij poezd : elektrihcka]; N = Masc, Fem, Neut, Plur; A == Msyl St. Chng, St Trunc partially or wholly, X, [motornaha lodka : motorka; anatomihceskij teatr : anatomihcka], except if Msyl St would result; A-St Trunc [-telqn-] and add [-l-], X, [zahzigatelqnaha bomba : zahzigalka]. Forms Fem N, occ An, X, [texnihceskaha sluhzakhaha : texnihcka]. Prod, more Prod than [-ik] in Incl. Stl, Col.
- 29 [-telqn(yj)]: Affxd, replaced by [-l-] before [-k(a)] in Incl, X, [zahzigatelqnaha bomba : zahzigalka].
- 29 [-l-]: Affxd A-St where [-telqn-] Trunc before [-k(a)] in Incl. Chng, St Trunc [-telqn-] before [-l-k(a)] in Incl, X, [zahzigatelqnaha bomba : zahzigalka].

- 29 [-ik]: Affxd A-St of AN Comb; sel Incl mn of Fem N, X, [maxovoe koleso : maxovik]; A-St not in [-sk-]. less Prod than [-k(a)] in Incl; esp Prod with A in [-ov-]. Forms Masc N, occ An, esp with A-St in [-ov-], X, [kadrovoy rabohcij : kadrovik].
- 29 [-sk(ij)]: Forms, does not Form St for Der in Incl before [-ik].
- 29 [-ov-]: Forms, often Forms St for Der of Masc N before [-ik] in Incl.
- 30 [-uhsk(a)]: Affxd A-St of AN Comb; Incl mn of N. not Prod. X, [teploe pomehhenie : tepluhska].
- 30 [-nic(a)]: Affxd A-St of AN Comb; Incl mn of N. not Prod. X, [poloskatelqnaha hcahska : poloskatelqnica].
- 30 [-ahc]: Affxd A-St of AN Comb; Incl mn of N. not Prod. X, [strogij vygovor : strogahc].
- 30 [-ov(oj)]: Affxd N-St. Forms An Subzd A, X, [gorn : gornovoj].
- 33 [#]: Affxd A of AN Comb; Incl mn of N. limited Prod. X, [stolihcnaa vodka : stolica].

8.4. Abstractors' Format with Abbreviations Resolved

Following is the same excerpt with the abbreviations resolved:

[Hanko-Trinickaha, N.A., "Processy vkluhceniha v leksike i slovoobrazovanii," Razvitie grammatiki i leksiki sovremennogo russkogo hazyka], Moscow, 1964, pp. 18-35.

- 25 [-sha]: Affixed to transitive verbs. Means 'inanimate, reflexive, mutual or generalized object'.
- 27f [-k(a)]: Affixed to the adjectival stem of an adjective-noun combination; it includes the meaning of the noun, for example [helektrihceskij poezd : elektrihcka];

the noun can be masculine, feminine, neuter or plural; the adjective cannot have monosyllabic stem. Change, the stem is truncated partially or wholly, for example [motornaha lodka : motorka; anatomihceskij teatr : anatomihcka], except if a monosyllabic stem would result; adjective stems truncate [-telqn-] and add [-l-], for example [zahzigatelqnaha bomba : zahzigalka]. Forms feminine nouns which are occasionally animate, for example [texnihceskaha sluhzakhaha : texnihcka]. It is productive, more productive than [-ik] in inclusion. Style, colloquial.

- 29 [-telqn(yj)]: Affixed, replaced by [-l-] before [-k(a)] in the process of inclusion, for example [zahzigatelqnaha bomba : zahzigalka].
- 29 [-l-]: Affixed to adjectival stems where [-telqn-] is truncated before [-k(a)] in the process of inclusion. Change, stems truncate [-telqn-] before [-l-k(a)] in inclusion, for example [zahzigatelqnaha bomba : zahzigalka].
- 29 [-ik]: Affixed to the adjectival stem of an adjective-noun combination; it seldom includes the meaning of feminine nouns, for example [maxovoe koleso : maxovik]; the adjectival stem cannot end in [-sk-]. It is less productive than [-k(a)] in inclusion; it is especially productive with adjectival stems in [-ov-]. Forms masculine nouns which are occasionally animate, especially with adjective stems in [-ov-], for example [kadrovoy rabochij : kadrovik].
- 29 [-sk(ij)]: Forms, does not form stems for derivation in inclusion before [-ik].
- 29 [-ov-]: Forms, often forms stems for derivation of masculine nouns before [-ik] in inclusion.
- 30 [-uhsk(a)]: Affixed to the adjectival stem of an adjective-noun combination; includes the meaning of the noun. It is not productive. For example, [teploe pomehhenie : tepluhska].

- 30 [nic(a)]: Affixed to the adjectival stem of an adjective-noun combination; includes the meaning of the noun. It is not productive. For example [poloskatelqnaha hcahska : poloskatelqnica].
- 30 [-ahc]: Affixed to the adjectival stem of an adjective-noun combination; it includes the meaning of the noun. It is not productive. For example, [strogij vygovor : strogahc].
- 30 [-ov(oj)]: Affixed to noun stems; forms animate substantivized adjectives, for example [gorn : gornevoj].
- 33 [#]: Affixed to the adjective of an adjective-noun combination; includes the meaning of the noun. Its productivity is limited. For example, [stolihcnaha vodka : stolica].

JMENT CONTROL DATA

GINATING ACTIVITY

THE RAND CORPORATION

2a REPORT SECURITY CLASSIFICATION

UNCLASSIFIED

2b GROUP

REPORT TITLE

A DEEP INDEX OF DERIVATIONAL MORPHOLOGY

AUTHOR(S) (Last name, first name, initial)

Worth, Dean S.

REPORT DATE

September 1966

6a TOTAL No OF PAGES

37

6b No OF REFS

CONTRACT OR GRANT No.

AF 49(638)-1700

8 ORIGINATOR'S REPORT No

RM-5143-PR

AVAILABILITY/LIMITATION NOTICES

DDC-1

9a SPONSORING AGENCY

United States Air Force

Project RAND

ABSTRACT

A discussion of a projected "deep index" to the extant literature on Russian derivational morphology. The Memorandum describes the type of literature being abstracted, the classification of information into discrete categories for retrieval, the steps taken to prepare data for conversion to machine format, and the transliteration and abbreviation conventions used in the index. An example of Russian scholarly text is appended, together with the format in which this text appears after processing.

11 KEY WORDS

Information storage and
retrieval
Linguistics
Russian language